

SECURE SPEECH

FIELD OF THE INVENTION

The present invention relates to apparatus, methods and computer readable media for providing an environment with secure speech capabilities. More particularly, the invention relates to providing such an environment by using noise canceling and noise covering techniques. The present invention can be deployed on various types of portable devices including, but not limited to phones, PDAs, pagers, and portable computers.

BACKGROUND OF THE INVENTION

Most business travelers bring work with them on trips to maximize the useful time away from the office, and to give them something to do on an airplane or train. They usually bring along a portable computer of some type so that they can edit or create documents and presentations, or write software. Most of these travelers also carry a cellular phone, and some carry a dictating device such as a tape recorder or digital voice recorder. The recording is used to make notations or to dictate documents for a secretary to transcribe when the traveler returns to the office.

Dictating information in an airplane can be distracting to the other passengers and can also be uncomfortable for the speaker as some of the dictated information may be sensitive. No one wants to sit next

1 to someone that is continually dictating, especially on
2 a late "red-eye" flight. It would be desirable for
3 there to be no noise or distraction while speaking or
4 dictating.

5 There is a considerable amount of prior art in the
6 fields of noise cancellation and echo cancellation.
7 Some of these methods employ algorithms and
8 technologies to remove background noise, echo effects,
9 and other undesirable content from the desired audio.
10 Although these systems have been developed to remove
11 the noise component from speech, almost none of them
12 have been used to cancel the entire speech audio
13 content.

14 One approach that does attempt to cancel the
15 speech content is set forth in United States Patent No.
16 5,526,421 to Berger et al. In Berger et al., a voice
17 transmission system, such as a telephone, includes a
18 microphone in combination with active sound
19 cancellation means comprising a speaker and a signal
20 processor for generating a mirror-image wave form with
21 respect to the signal generated when speaking into the
22 microphone. The mirror-image wave form is used to
23 activate the speaker, thereby canceling the user's
24 voice.

25 Unfortunately this approach does not necessarily
26 provide the user with complete privacy, since the
27 cancellation is never perfect and does not prevent the
28 user's voice from being overheard by others. This is

1 because the loudspeaker or other sound source which
2 provides the canceling audio is not identical to the
3 original source, and is at a location which is at least
4 slightly displaced therefrom. While the speaker's
5 voice may be at least partially canceled, it is
6 virtually impossible to match frequency, phase and
7 amplitude so as to have complete cancellation in all
8 regions of the space surrounding the speaker. Under
9 these circumstances, there is a large probability that
10 someone in the vicinity of the speaker will be able to
11 hear and understand what the speaker is saying. The
12 result is a lack of complete privacy, and the very
13 purpose of the system is defeated.

14 SUMMARY OF THE INVENTION

15 It is therefore an aspect of the present invention
16 to provide a system and a method for canceling a
17 speaker's voice and masking it, so that complete
18 privacy is assured.

19 It is another aspect of the invention to provide a
20 masking signal, and apparatus for generating it, that
21 is related to the sounds that are to be masked.

22 In accordance with the invention, sound is
23 converted into an electrical signal by a microphone and
24 the electrical signal is analyzed. The electrical
25 signal also is converted into another signal that is
26 the exact opposite of the incoming electrical signal.
27 This resulting signal is then digitally modulated,

1 amplified, and sent back by a speaker or speakers
2 mounted in the device, canceling out the audible
3 portion of the sound.

4 The invention may be implemented by a specialized
5 electronic circuit and associated software algorithms
6 embodied by the specialized circuitry and other
7 programs installed on the device. The circuit may
8 employ techniques and methods normally found in the
9 area of noise reduction. In accordance with the
10 invention, as in some noise reduction applications, the
11 noise component of a particular signal is detected,
12 often using a statistical characterization. The noise
13 component is then removed from the signal, either by
14 introducing an identical noise component having a phase
15 opposite to that of the original, or by some other
16 means. The signal components that are 180 degrees out
17 of phase with each other cancel each other out,
18 removing the noise component.

19 The present invention provides a similar type of
20 method, but instead of applying it to the noise
21 component, the method is applied to the main voice
22 component. In this case, the person speaking into a
23 microphone causes a signal to be produced. This signal
24 is immediately analyzed, then inverted, and
25 retransmitted using speakers built into the device.
26 The speakers are designed to create an acoustic field
27 that is directed in the same way as the person's voice,
28 except that the field is minimized in the direction of
29 the microphone. Thus the microphone receives the

1 person's voice alone; all other listeners receive both
2 the person's voice and the output of the speakers.
3 Since these are equal in amplitude and opposite in
4 phase, they cancel, so that other listeners hear
5 nothing. The user is then free to dictate or speak into
6 a device without others being able to discern the
7 speaker's content.

8 The invention employs the phenomenon called
9 "acoustic masking" in recognizing that one sound, if
10 of sufficient amplitude, can make another sound
11 difficult or impossible to hear. Using this phenomenon,
12 the device generates a noise component that disguises
13 or hides the speaker's content. That noise is
14 subtracted from the voice component. The frequency of
15 the noise can also be shifted within the audible range
16 to discourage eavesdropping.

17 Thus, in accordance with the invention, a method
18 for masking speech, a system having components for
19 masking speech, and an article of manufacture having
20 computer readable program code thereon, the code
21 causing the computer to perform, the method are all
22 provided. The method comprises generating an
23 electrical signal representative of the speech, using
24 the electrical signal to provide an audio signal which
25 cancels the speech; and providing a speech masking
26 signal to mask any speech not canceled. The speech
27 masking signal is supplied by a babble generator. The
28 speech masking signal may be produced by rearranging
29 the speech so that it is not intelligible.

1 The speech masking signal may be provided by pitch
2 inverting the speech. The pitch may be inverted by low
3 pass filtering the electrical signal representative of
4 the speech to provide a low pass filtered signal;
5 mixing the low pass filtered signal with a carrier
6 frequency to produce an output signal including the
7 carrier frequency and two side bands; and low pass
8 filtering the output signal to produce the masking
9 signal.

10 The speech masking signal may also be produced
11 reading out digital representations of signals
12 disruptive to the understanding of speech from a
13 memory; and converting the digital representations to
14 the speech masking signal. The digital representations
15 may be in the form of one of pulse code modulation and
16 adaptive pulse code modulation, and may be
17 representative of multiple persons speaking
18 simultaneously. The speech masking signal may be
19 produced by sampling portions of the speech; and
20 providing the portions in reverse order. At least one
21 of amplitude and frequency of the masking signal may be
22 changed at intervals. The intervals may be fixed,
23 variable or random.

24
25 The speech masking signal may be produced by
26 sampling portions of the speech; and providing the
27 portions with a predetermined time delay. In this case
28 also, the speech masking signal may be produced by
29 sampling portions of the speech; and providing the

1 portions in reverse order and with a predetermined time
2 delay. Again, at least one of amplitude and frequency
3 of the masking signal may be changed at fixed, variable
4 or random intervals.

5
6 The speech masking signal may be produced by sampling
7 portions of the speech; and providing the portions in a
8 random manner. The portions may be provided so as to
9 be random relative to amplitude and frequency of the
10 speech.

11 BRIEF DESCRIPTION OF THE DRAWINGS

12 These and other aspects, features, and advantages
13 of the present invention will become apparent upon
14 further consideration of the following detailed
15 description of the invention when read in conjunction
16 with the drawing figures, in which:

17 Fig. 1 is a block diagram of a system in
18 accordance with the invention;

19 Fig. 2 is a block diagram of a first embodiment of
20 a babble generator which may be used in the system of
21 Fig. 1; and

22 Fig. 3 is a block diagram of a second embodiment
23 of a babble generator which may be used in the system
24 of Fig. 1.

1 Fig. 4 is a block diagram of a third embodiment of
2 a babble generator which may be used in the system of
3 Fig. 1.

4 DESCRIPTION OF THE INVENTION

5 One form of preferred embodiment is a circuit
6 consisting of an amplifier, a analog-to-digital
7 converter, an inverter, a digital-to-analog converter,
8 a power amplifier and two or more speakers. The
9 invention may then be embodied as a computer program
10 (or as computer readable program code stored on a
11 computer readable storage medium) in a computing device
12 that inverts the phase of an incoming signal and uses
13 that inversion to cancel or disguise the sound of the
14 speaker's voice, thus preventing other individuals in
15 the general vicinity of the speaker from hearing
16 clearly what the user is saying. This system can also
17 be implemented using a microphone coupled to an
18 analog-to-digital converter and then to the circuitry
19 described in Fig. 1.

20 It will be understood that the system of Fig. 1
21 may be implemented as a stand alone device used with,
22 for example, a telephone, so that the user may have a
23 private telephone conversation with someone at the
24 other end of the telephone without fear of a private
25 conversation being overheard by someone else. However,
26 as noted above, the system and method of the invention
27 may also be used with a device for recording dictation
28 or with a computer. It may be provided as a computer
29 peripheral device. Alternatively, the system may be

1 implemented in a computer, using the computer's sound
2 card for certain functions, and a computer program to
3 perform other functions. The computer may be used to
4 implement the invention, while being connected to a
5 telephone system to allow a private conversation to
6 take place. In this case, as noted above, the system
7 according to the invention is implemented by the
8 computer, rather than a separate stand alone piece of
9 equipment. In addition, the invention may be used when
10 a computer is performing speech recognition tasks,
11 including speech directed navigation and voice
12 transcription.

13 Referring specifically to Fig. 1, when the user
14 speaks into a microphone 101, the signal 100 is
15 generated by the microphone and its associated support
16 circuitry, including an analog-to digital converter or
17 ADC 111. The signal is routed to a first input of a
18 subtractor 115 to subtract any component of the
19 canceling audio from the normal incoming audio, which
20 is processed by an inverting amplifier Amp/Inverter
21 105, and applied to a second input of subtractor 115.
22 The resultant audio is routed to the normal audio
23 handling features of a device memory 106. This may be a
24 recording head, memory, disk, or any device used to
25 record, process, or play the input audio signal 100.
26 The input signal 100 is also routed from ADC 111 to
27 amplifier/inverter 105 where the signal is processed
28 and inverted for output. The inverted output signal is
29 then routed to a digital-to-analog converter DAC/Delay
30 112, then to a power amplifier 113 and then to a

1 speaker or speakers 107 where it is output from the
2 device system. The output signal, being 180 degrees
3 out of phase with the input signal, causes the
4 speaker's voice to be canceled in the general proximity
5 of the user. This assists in prohibiting others from
6 listening in on the user's conversation or dictation,
7 but as noted above, is not always sufficient.

8 The digital signal at Amp/Inverter 105 may be
9 optionally time shifted a small amount to compensate
10 for latencies in components and software, and is then
11 converted to analog at DAC/Delay 112. The signal is
12 then routed to the cancellation speaker 107.

13 The system in accordance with the present
14 invention also contains a babble generator 110 (Fig.
15 2), 110A (Fig. 3), or 110B (Fig. 4) which may generate
16 human-like "babble" by, for example, sampling the
17 digital input and sending a random set of digital
18 phonemes which are converted to analog and output
19 through the cancellation speaker. The phonemes may be
20 generated in several ways, as described below.

21 Referring to Fig. 2, a babble generator 110
22 receives input from input signal 100 of Fig. 1 along a
23 signal line 115. This form of babble generator is
24 similar to that used for secure analog telephones. It
25 comprises a pitch shifter followed by a filter. The
26 shifter includes a low-pass filter 200, which
27 band-limits the speech signal so that it does not have
28 any frequency components exceeding, for example, 3 kHz.

1 This band-limited signal is input to a multiplier 201
2 having as another input a 3 kHz signal from a local
3 oscillator 202. The output of multiplier 201 is a
4 signal on a line 203 containing the input signal 100
5 shifted to the 3-6 kHz range, the 3 kHz carrier, and a
6 signal that is the input signal 100, pitch-inverted.
7 That is, if the input signal contains a 1 kHz
8 component, the pitch-inverted signal contains a 2 kHz
9 component. If the input signal contains a 2.5 kHz
10 component the pitch-inverted signal contains a 500 Hz
11 component. The output of the multiplier is input into a
12 low-pass filter 204, which is preferably identical to
13 low-pass filter 200. The output of that low-pass filter
14 is a signal on a signal line 205, which is a
15 pitch-inverted version of input signal 100, and is
16 mixed with the output of DAC/delay 112 (Fig. 1) to form
17 the input to power amplifier 113.

18 Referring to Fig. 3, a babble generator 110A uses
19 phoneme recognition and anti-phoneme generation. This
20 approach relies on the fact that speech can be broken
21 down into several primary types of phonemes called
22 stops, fricatives, affricates, nasals, liquids, and
23 glides. Each category describes a particular sound
24 using a consonant or group of consonants. For example,
25 the sound made with "ch" is fricative or affricate,
26 while the sound made by the "y" is referred to a glide.

27 In Fig. 3, using a rule-based phoneme recognizer,
28 the incoming text is sampled at the user's input after
29 low pass filtering and converted to phonemes in a
30 phoneme recognizer 210. The conversion of analog voice

1 signals to phonemes is a well understood technology for
2 which there are several different methods available. In
3 the present invention, these methods may be
4 encapsulated in an electrical component or chip, or
5 performed in software. When the phonemes are detected,
6 a shifting or delay may be applied in a phoneme data
7 shifter/delay 212. The output of shifter/delay 212 is
8 applied to a phoneme generator 214. Generator 214
9 creates a negation to the phoneme, creating a
10 cancelling sound or what may be referred to as an
11 "anti-phoneme". These "anti-phonemes" are then used as
12 output to be applied to the input of amplifier 113,
13 thus producing a babble content that has the capacity
14 to render the user's speech unintelligible. A phoneme
15 order generator 216 may also be utilised, as described
16 below.

17 In another form of the present invention, which
18 may utilize a portion of the approach of Fig. 2, the
19 phonemes detected by the phoneme recognizer 210 and
20 subsequently generated by the phoneme generator 214 are
21 shifted up or down in frequency and or amplitude, and
22 may also be delayed a short amount of time prior to
23 being output the speaker. The techniques for shifting
24 frequency or delaying a signal are well understood.

25 In yet another form of the present invention, the
26 phonemes detected by the phoneme recognizer 210 and
27 subsequently generated by the phoneme generator 214 are
28 randomly replaced or modified using a form of a phoneme
29 order generator 216 that selects at random, or in a

1 reverse order, a phoneme or phonemes, respectively, to
2 substitute for existing phonemes, prior to output to
3 the speaker. If operating in a random mode, the order
4 of the generated phonemes is a random variation of the
5 order of the recognized phonemes. If operating in a
6 reverse order mode, then during a given fixed or
7 random, but relatively short time interval, phonemes
8 are delayed and played back in reverse order. The
9 amplitudes of the phonemes may also be varied in a
10 random manner, at fixed or random intervals. These
11 approaches all serve to further assist in rendering the
12 speech unintelligible.

13 The embodiments of the babble generators 110 and
14 110A of Fig. 2 and Fig. 3, respectively, which derive
15 input from the speaker, have a major advantage in that,
16 subject to a very short delay, there are no masking
17 sounds when the speaker is not speaking. This permits
18 the area to be quiet when speech does not need need to
19 be masked. With suitable modifications, such as the
20 addition of a speech level detector and a switching
21 circuit, having a comparator responsive to the speech
22 level, the embodiment of Fig. 4 described below may be
23 modified to operate in a similar manner.

24 Referring to Fig. 4, a babble generator 110 is not
25 driven by input signal 100. Block 300 comprises a
26 read-only serially-accessible memory containing a
27 digitized version of multiple speakers' speech, as
28 might be recorded at a cocktail party, for example. The
29 output 301 of block 300 is in some convenient

1 representation such as Pulse-Code Modulation (PCM) or
2 Adaptive Digital Pulse-Code Modulation (ADPCM),
3 well-known in the art. This representation is input
4 into a digital-to-analog converter 302 that can decode
5 the representation to reconstruct the speech as
6 originally digitized. Any type of speech can be used,
7 as can other signals not related to speech but known to
8 be disruptive to the understanding of speech. The
9 output of block 302 is signal 303 which is the required
10 babble output, and is mixed with the output of
11 DAC/delay 112 to form the input to power amplifier 113.

12 Variations described for the present invention
13 can be realized in any combination desirable for each
14 particular application. Thus particular limitations,
15 and/or embodiment enhancements described herein, which
16 may have particular advantages to the particular
17 application need not be used for all applications.
18 Also, it should be realized that not all limitations
19 need be implemented in methods, systems and/or
20 apparatus including one or more concepts of the present
21 invention.

22 The present invention can be realized in
23 hardware, software, or a combination of hardware
24 and software. Any kind of computer system, or other
25 apparatus adapted for carrying out the methods and/or
26 functions described herein, is suitable. A typical
27 combination of hardware and software could be a
28 general purpose computer system with a computer program
29 that, when being loaded and executed, controls the

1 computer system such that it carries out the methods
2 described herein. The present invention can also be
3 embedded in a computer program product, which comprises
4 all the features enabling the implementation of the
5 methods described herein, and which, when loaded in a
6 computer system, is able to carry out these methods.

7 Computer program means or computer program in
8 the present context include any expression, in any
9 language, code or notation, of a set of instructions
10 intended to cause a system having an information
11 processing capability to perform a particular function
12 either directly or after conversion to another
13 language, code or notation, and/or reproduction in
14 a different material form.

15 Thus the invention includes an article of
16 manufacture which comprises a computer usable medium
17 having computer readable program code means embodied
18 therein for causing a function described above.
19 The computer readable program code means in the
20 article of manufacture comprises computer readable
21 program code means for causing a computer to effect the
22 steps of a method of this invention. Similarly, the
23 present invention may be implemented as a computer
24 program product comprising a computer usable medium
25 having computer readable program code means embodied
26 therein for causing a function described above. The
27 computer readable program code means in the computer
28 program product comprising computer readable program

1 code means for causing a computer to effect one or
2 more functions of this invention. Furthermore, the
3 present invention may be implemented as a program
4 storage device readable by machine, tangibly embodying
5 a program of instructions executable by the machine to
6 perform method steps for causing one or more functions
7 of this invention.

8 It is noted that the foregoing has outlined some
9 of the more pertinent objects and embodiments of the
10 present invention. The concepts of this invention may
11 be used for many applications. Thus, although the
12 description is made for particular arrangements and
13 methods, the intent and concept of the invention is
14 suitable and applicable to other arrangements and
15 applications. It will be clear to those skilled in the
16 art that other modifications to the disclosed
17 embodiments can be effected without departing from the
18 spirit and scope of the invention. The described
19 embodiments ought to be construed to be merely
20 illustrative of some of the more prominent features and
21 applications of the invention. Other beneficial
22 results can be realized by applying the disclosed
23 invention in a different manner or modifying the
24 invention in ways known to those familiar with the art.
25 Thus, it should be understood that the embodiments has
26 been provided as an example and not as a limitation.
27 The scope of the invention is defined by the appended
28 claims.